FABRIC DEFECT DETECTION BASED ON FASTER RCNN WITH CBAM

Yuan He, Han-Dong Zhang, Xin-Yue Huang
and Francis Eng Hock Tay

Department of Mechanical Engineering,
National University of Singapore, Singapore

ABSTRACT

In the production process of fabric, defect detection plays an important role in the control of product quality. Consider that traditional manual fabric defect detection method are time-consuming and inaccuracy, utilizing computer vision technology to automatically detect fabric defects can better fulfill the manufacture requirement. In this project, we improved Faster RCNN with convolutional block attention module (CBAM) to detect fabric defects. Attention module is introduced from graph neural network, it can infer the attention map from the intermediate feature map and multiply the attention map to adaptively refine the feature. This method improve the performance of classification and detection without increase the computation-consuming. The experiment results show that Faster RCNN with attention module can efficient improve the classification accuracy.

KEYWORDS

Fabric defects detection, Faster RCNN, Convolutional block attention module, Deep learning.

1. INTRODUCTION

In order to produce highquality garments, it is an important step to apply a defect detection link in the process of fabric manufacturing to ensure the quality. Defect detecting is the process to find out and locate defects on the surface of fabric. Finding out defects on fabric also improves the efficiency of manufacturing process by abandoning unqualified intermediate products. Traditionally, manual inspection which carried out on wooden board is the only method to assure the quality of textile. Sometimes workers also do fine defects detection with the help of equipment like magnifiers and microscopes. Manual defect detection can do prompt correction of small defects. However, error may occur due to fatigue, and small defects are usually undetected [1].

Since fabric defect detection has a great effect on the quality control of textile manufacture and the conventional manual inspection method does not suit the requirement of developed automated manufacture, automatic fabric defects detection becomes a natural way to improve fabric quality and lower labor cost. Fortunately, with the development of deep learning technology and the progress of computer vision technology, a new automated fabric detection method which can replace manual inspection appears. By applying computer vision and machine learning technology, automated visual inspection is widely used to detect the surface defects of machined parts and components. According to the research of Rajalingappaa Shanmugamani [2] published in 2015, visual inspection method can provide rapid quantitative assessment and improve quality and productivity.

David C. Wyld et al. (Eds): CONEDU, NBIoT, MLCL, ISPR, CSITA - 2021
pp. 99-107, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.110608

Two defects detection algorithms are compared in this project, the Faster RCNN and Faster RCNN with convolutional block attention module (CBAM). The difference between these two algorithms is that the backbone net is different. The backbone net for Faster RCNN is Resnet-50[21], which is a 50 layers deep neural network and used for feature extraction of defects,classification and regression. CBAM will combine with Resnet-50 to improve the performance. Both algorithm detector is Faster RCNN. By comparing the result of two detect algorithms, the role of attention module will be revealed, and the effect of the faster R-CNN is going to be shown.

2. RELATED WORK

At present, textile defects detection approaches can be simply divided into spectral approach and learning approach. Gabor filters provide the optimal joint position in spatial and frequency domain [3], it becomes the most popular approach in spectral-based method. The initial application of Gabor filter is to build a filter bank with numerous sets of filters, which is predetermined the parameters in frequency and orientation [4]. In [5], Shu calculates the frequency and direction data obtained from 16 Gabor filters convolution with 4 different angles and scales to detect fabric defects. This method accuracy will be affected by the computationally intensive and the frequency plane coverage. Bodnarova [6] utilizes optimal filters to reduce the amount of filters, the computation time is greatly reduced resulting in an increased speed of detection. However, the correct choice of optimal filter is difficult and crucial. Tong [7] has developed composite differential evolution (CoDE) to optimize the parameters of Gabor filters, and get high performance in limited samples. LI [8] integrated Gabor filter and Gaussian mixture model to inspect simple texture defects, the classification accuracy from 360 images of 9 different defect types reach 87%.

Neural network has advantages in feature extraction, segmentation and optimization tasks of fabric defects detection area [9,10]. In 2001, Stojanovic[11] proposed a three-layer back-propagation neural network for low-cost fabric real-time detection, and the accuracy achieved 86%. Kumar [12] combines forward neural network with Principal Component Analysis (PCA) for faster detection. In [13], Kuoproposes a three-layers back-propagation neural network to detect white fabric defects. This model a high dimensional system by non-linear regression algorithm, achieve 91.88% recognition accuracy for 160 simple defects image. Asimilar architecture of network is proposed in [14], the detection accuracy of holes and oil stain of twill fabric reach 91% and 100%. However, the amount of sample is limit and the reliability is unknown. Semnani and Vadood [15] develop an intelligent model based on artificial neural network to estimate the appearance of knitted fabrics.

With the development of deep learning, such as R-CNN [16], fast R-CNN [17], SSD [18] and YOLO [19], fabric defects automatically detection has more potential possibilities. R-CNN is the abbreviation of Region-based Convolutional Neural Networks and it is put forward by Girshick and his team since 2014 [16]. This method mainly includes two steps. In the first step, R-CNN applies a controllable amount of bounding boxesto select candidates.ThenR-CNN independently extracts features from each candidate to classification. To improve the performance of R-CNN and increase its operation speed, Girshick (2015) [17] reconstructed the network architecture. Girshick combines three independent CNN into one joint trained framework andshare union parameters. Fast R-CNN integrates the scattered feature vectors into a feature matrix shared by all the region proposals. The feature matrix is realized by the forward propagate of CNN on the entire image, and then it is separated to train the classifier and regressor to classification and point out their locations.

3. METHOD

3.1. Date augmentation

In this paper, we collected 185 fabric defect images in three different typesfrom textile factory. They are broken hole, fly yarn and drop needle. Since the database is too small, we adapt random real-time data augmentation to increase the amount of image, we flip the image in horizontal and vertical direction, and distort the image.

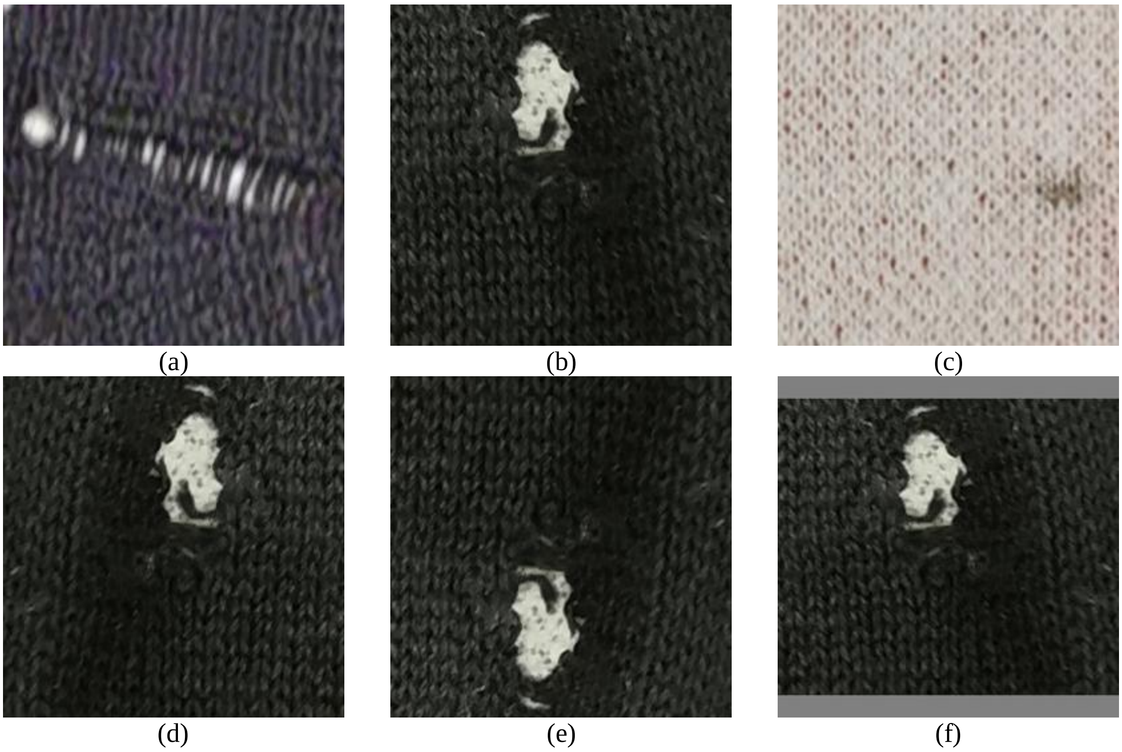

Fig 1. The examples of fabric defects image and augmentation results. (a) is drop needle. (b) is broken hole. (c) is fly yarn. (d) and (e) are (b) flipped in horizonal and vertical direction, (f) is (b) distorted and resized in standard size.

3.2. Faster RCNN

Faster R-CNN does the job of object detection in this project. It is developed based on R-CNN and Fast R-CNN technology and was proposed by Girshick and his team in 2015 [20].Faster R-CNN integrates the step of creating boundary boxes into CNN model. The overall frame of Faster R-CNN is shown in Figure 2. Faster RCNN contains four main parts, which are convolution layers, region proposal network, region of interest (RoI) pooling and classification.

In this paper, we use ResNet-50 [21] as the backbone net. Firstly, the input fabric defects image feature map is extracted by convolutional layer with ReLu activation function and pooling layers.Region proposal network (RPN) is used to generate proposals with feature matrix. We use boundary conditions and non-maximum suppression [22] to select the appropriate anchor on feature map. The output of regression layer indicates the coordinate position of fabric defect. RoI pools collect region proposal and feature maps, the bounding box regression provides the final exact target box position.

CBAM (Convolutional Block Attention Module) [23] is an efficient improvement algorithm in detection presented by Sanghyun Woo, Jongchan Park and their team in 2018.The structure of these two modulesshows in Figure 3.The function of the entire attention module can be expressed by the following two equations. The first equation represents the function of channel attention module and the second equation shows the function of spatial attention module.

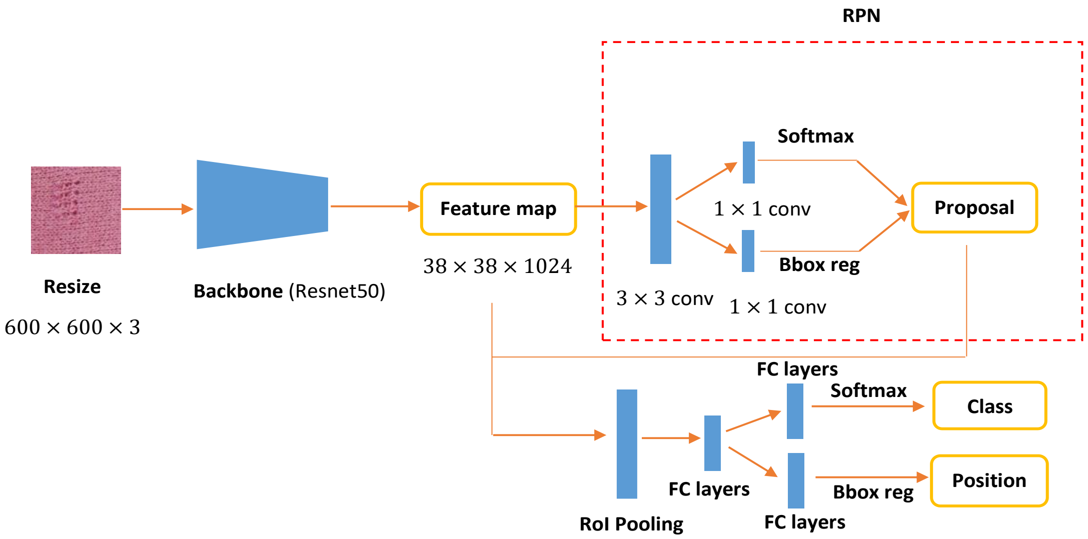

Fig 2. An illustration of Faster R-CNN model

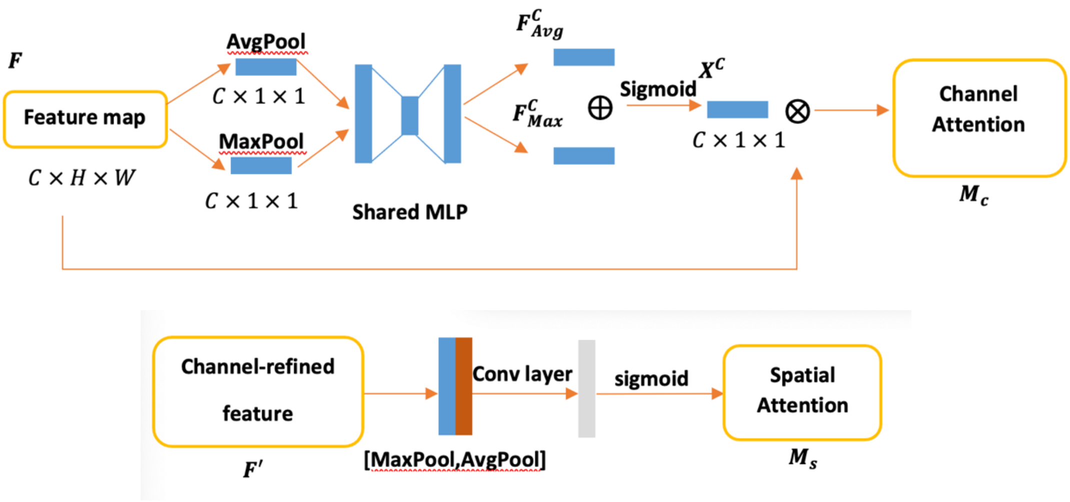

Fig 3. Diagram of each attention sub-modules

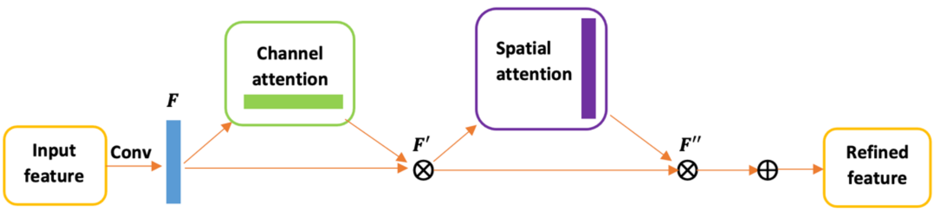

Fig 4. CBAM integrated with a ResBlock in ResNet

$$\begin{aligned} F' &= M_c(F) \oplus F \\ F'' &= M_s(F') \oplus F' \end{aligned} \tag{1}$$

$\oplus$ represents element-wise multiplication. M_c represents the operation of attention extraction on the channel dimension, and M_s represents the operation of attention extraction on the spatial dimension. $F \in \mathbb{R}^{H\times W\times C}$ is the intermediate feature map, F' is the product of feature map after channel attention module process, F'' represents the final output feature map after channel attention and spatial attention. H means the height of input feature map, W is the width of input feature map. C is the channel number of input feature map.

Table1. ResNet-50 Architecture

Layer Name	**Output Size**	**ResNet-50**
conv1	$112 \times 112 \times 64$	7×7, 64, stride 2
conv2_x	$56 \times 156 \times 64$	3×3 max pool, stride 2 $\begin{bmatrix} 1\times 1, 64 \\ 3\times 3, 64 \\ 1\times 1, 256 \end{bmatrix} \times 3$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 1\times 1, 128 \\ 3\times 3, 128 \\ 1\times 1, 512 \end{bmatrix} \times 4$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 1\times 1, 256 \\ 3\times 3, 256 \\ 1\times 1, 1024 \end{bmatrix} \times 23$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 1\times 1, 512 \\ 3\times 3, 512 \\ 1\times 1, 2048 \end{bmatrix} \times 3$
average pool	$1 \times 1 \times 512$	
softmax	1000	

As shown in Figure 3, the input feature map suffersglobal average-pooling F^c_{avg} and globalmaxi-pooling F^c_{max} in channel attention module. It is worth to note that the global average pooling and global maximum pooling are used in parallel, whichcan minimize the information loss during the pooling process. F^c_{avg} and F^c_{max} forward to a shared network which composed of multi-layer perceptron with one-hidden layer, to produce channel attention map M_c. It can be represented as following equation (2):

$$\begin{aligned} M_c(F) &= \sigma\Big(MLP\big(AvgPool(F)\big) + MLP\big(MaxPool(F)\big)\Big) \\ &= \sigma\big(W_1(W_0(F^c_{avg})) + W_1(W_0(F^c_{max}))\big) \end{aligned} \tag{2}$$

Where σ presents sigmoid function, $W_0 \in \mathbb{R}^{C/r\times C}$,$W_1 \in \mathbb{R}^{C\times C/r}$is the MLP layers wight, r is the reduction rate. First layer has ReLu activation function.

Different from channel attention module, spatial attention module applies a convolutional layer to generate a spatial attention map to concatenate max-pooling and average-pooling. It can be represented as following equation (2):

$$M_c(F) = \sigma\Big(f^{7\times 7}\big(AvgPool(F)\big); MLP\big(MaxPool(F)\big)\Big) \tag{3}$$

$$= \sigma\left(f^{7\times7}\left[F^{s}_{avg}; F^{s}_{max}\right]\right)$$

Where σ presents sigmoid function, $f^{7\times7}$ represents a convolution operation with 7×7 size filter. In this paper, we use sequential arrangement attention modules.

4. EXPERIMENT AND RESULT

In this paper, we used 90 percent samples in the database as training dataset. For one experiment, 50 epochs are run with a base learning rate of 0.001 and 0.0001 with another 50 epochs. In both of the two iteration stages, the learning rate will decay to 92% of the original in each iteration. Batch size is 1 and we carry 2 experiments for each method separately. Totally, we have taken 40,000 iterations on each method. The test environment is a HP desktop with an Inter(R) Core(TM) i5-4200 3.3 GHZ CPU, the simulation software is python2.7.

There are 185 images in the database, 10% of which is test dataset. Limited by the sample size, only 3 types of defects are involved in training and testing. All of the three types of defects can be detected under both the Faster R-CNN with a modified backbone net and the normal one as shown in the Fig 3. we just show some examples of the detection results. It can be seen that both these two methods can detect the fabric defects accurately. For the two figures in each group, the left one is the result of the normal Faster R-CNN using Resnet 50 as backbone; the right one is the result of the Faster R-CNN with a modified backbone using Resnet50 and CBAM. Both of the two experiments use Faster R-CNN as the detector. It is clear that using Faster R-CNN with CBAM leads a higher confidence of the detected defect object compared with the normal one. After adapting CBAM, the confidence is 2%-3% higher than before. Therefore, the detection model can be regarded more reliable.

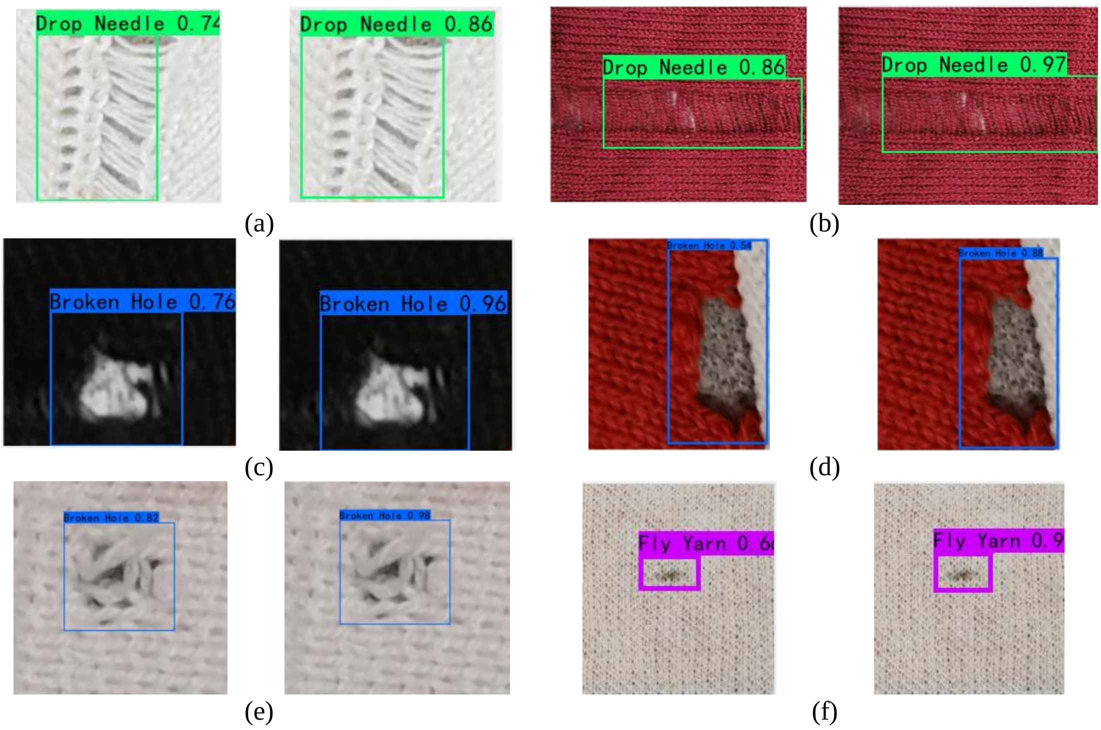

Fig 3. Results display and comparison between the different backbones: Resnet50 (left figure) and Resnet50 + CBAM (right figure). (a), (b) drop needle, (c), (d), (e)broken hole, (f) fly yarn.

The comparison and performance of the two methods are shown in the follow table separately. Average precision (AP) of each class and mean average precision (mAP) on the test dataset are used to evaluate the detection effect. The confidence threshold in mAP is 0.01. Training with CBAM can improve the mAP of fabric defects detection, it will become nearly 1% higher than before.

Table. 2 Object detection AP (%) of each class and mAP (%) on the test dataset.

Backbone	**Detector**	**AP of broken hole**	**AP of drop needle**	**AP of fly yarn**	**mAP**
ResNet-50	Faster R-CNN	93.26	61.03	55.50	69.93
ResNet-50+CBAM	Faster R-CNN	95.31	61.89	55.48	70.89

Global recognizable ratio is the percentage of samples that can be recognized in all test samples. Correspondingly, we can describe the percentage of samples that can be recognized in a certain class in all samples of that class as class recognizable ratio. Broken hole is the most easily to be detected since its structure feature is obvious and simple. As for drop needle, it is not very easy to be detected since it has many kinds of different structures in the database. Fly yarn is most difficult to be detected since its sample size is small and it does not have special structures.

Table. 3 Global recognizable ratio (%) and class recognizable ratio (%) on the test dataset.

Backbone	Detector	Global ratio	Broken hole	Drop needle	Fly yarn
ResNet-50	Faster R-CNN	46	86	50	36
ResNet-50+CBAM	Faster R-CNN	62	100	56	67

In addition, to determine whether every recognized sample is classified correctly, the ratio of samples with correct classification in all samples that can be recognized is accuracy. It can also be divided into global accuracy and category accuracy, like the recognizable ratio. Adapting CBAM makes the global accuracy becomes 1% higher than before. For each class, the accuracy of drop needle is the highest, although it is not easy to be detected out, its accuracy is high. With CBAM, broken hole can reach 100% accuracy and accuracy of fly yarn is also improved.

We also test the detection speed of Faster R-CNN which is 1.10 seconds per image.

Table. 4 Global accuracy (%) and class accuracy (%) of each class on the test dataset.

Backbone	Detector	Global accuracy	Broken hole	Drop needle	Fly yarn
ResNet-50	Faster R-CNN	87	83	100	71
ResNet-50+CBAM	Faster R-CNN	88	100	100	83

5. CONCLUSIONS

In this paper, we compared Faster RCNN and CBAM performance in fabric defects detection area. The experiment result proved CBAM can get better accuracy and recognizable ratio in our fabric defect database. Dueto the difference in image quality and pattern complexity, some defects cannot be successfully detected. Collecting fruitfulness high-quality fabric defect and defect-free image, updating existing network model will be the focus of future research. The different training time between CBAM and Faster RCNN is negligible. However, compared with the traditional approaches, the training time is more time-consuming. Reduce the required timefor training the model, realize real-time detection is crucial to whether this technology can be applied in industrial production, which is also the focus of future research.

REFERENCES

[1] Ngan, H. Y., Pang, G. K., & Yung, N. H. (2011). Automated fabric defect detection—a review. *Image and vision computing*, *29*(7), 442-458.

[2] Shanmugamani, R., Sadique, M., &Ramamoorthy, B. (2015). Detection and classification of surface defects of gun barrels using computer vision and machine learning. *Measurement*, *60*, 222-230.

[3] Kumar, A., & Pang, G. K. (2002). Defect detection in textured materials using Gabor filters. IEEE Transactions on industry applications, 38(2), 425-440.

[4] Mak, K. L., & Peng, P. (2008). An automated inspection system for textile fabrics based on Gabor filters. *Robotics and Computer-Integrated Manufacturing*, *24*(3), 359-369.

[5] Shu, Y., & Tan, Z. (2004). Fabric defects automatic detection using Gabor filters. In *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No. 04EX788)*(Vol. 4, pp. 3378-3380). IEEE.

[6] Bodnarova, A., Bennamoun, M., & Latham, S. (2002). Optimal Gabor filters for textile flaw detection. *Pattern recognition*, *35*(12), 2973-2991.

[7] Tong, L., Wong, W. K., &Kwong, C. K. (2016). Differential evolution-based optimal Gabor filter model for fabric inspection. *Neurocomputing*, *173*, 1386-1401.

[8] Zhang, Y., Lu, Z., &Li, J. (2009). Fabric defect detection and classification using gabor filters and gaussian mixture model. In *Asian conference on computer vision* (pp. 635-644). Springer, Berlin, Heidelberg.

[9] Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(1), 4-37.

[10] Egmont-Petersen, M., de Ridder, D., & Handels, H. (2002). Image processing with neural networks—a review. *Pattern recognition*, *35*(10), 2279-2301.

[11] Stojanovic, R., Mitropulos, P., Koulamas, C., Karayiannis, Y., Koubias, S., & Papadopoulos, G. (2001). Real-time vision-based system for textile fabric inspection. *Real-Time Imaging*, *7*(6), 507-518.

[12] Kumar, A. (2003). Neural network based detection of local textile defects. *Pattern Recognition*, *36*(7), 1645-1659.

[13] Kuo, C. F. J., Lee, C. J., & Tsai, C. C. (2003). Using a neural network to identify fabric defects in dynamic cloth inspection. *Textile Research Journal*, *73*(3), 238-244.

[14] Yin, Y., Zhang, K., & Lu, W. (2009). Textile flaw classification by wavelet reconstruction and BP neural network. In *International Symposium on Neural Networks* (pp. 694-701). Springer, Berlin, Heidelberg.

[15] Semnani, D., & Vadood, M. (2010). Improvement of intelligent methods for evaluating the apparent quality of knitted fabrics. *Engineering Applications of Artificial Intelligence*, *23*(2), 217-221.

[16] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2015). Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *38*(1), 142-158.

[17] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

[18] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.

[19] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

[20] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*.

[21] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[22] Neubeck, A., & Van Gool, L. (2006, August). Efficient non-maximum suppression. In *18th International Conference on Pattern Recognition (ICPR'06)* (Vol. 3, pp. 850-855). IEEE

[23] Woo, S., Park, J., Lee, J. Y., &Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).

AUTHORS

Yuan He is a Ph.d. Candidate in National University of Singapore, Singapore. He received the B.Sc. degress in Northeastern University, P.R. China. His research interests include defects detection, computer vision and image processing.

Dr. Francis Tay is currently an Associate Professor with the Department of Mechanical Engineering, Faculty of Engineering, National University of Singapore. Dr. Tay is the Deputy Director (Industry) for the Centre of Intelligent Products and Manufacturing Systems, where he takes charge of research projects involving industry and the Centre.

Zhang Handong is a master candidate in National University of Singapore, Singapore. He received the B.S. in Harbin Institute of Technology, P.R. China. His research interests include defects detection and mechanical design.

Xin-Yue Huang is a Master Candidate in University of Singapore, Singapore. She received the B.Sc. degree in Sichuang University, P.R. China. Her research interests include computer vision and image processing.